植 物 分 类 与 资 源 学 报 2013, 35 (2): 158~164
Plant Diversity and Resources DOI: 10.7677/ynzwyj201312070

欧亚大陆分布的大花菟丝子叶绿体基因组插入缺失分析*

王朝波[1,2], 龚 洵[1**]
(1 中国科学院昆明植物研究所资源植物与生物技术重点实验室, 云南 昆明 650201;
2 中国科学院大学, 北京 100049)

摘要: 运用鸟枪法在 Illumina 测序仪上对亚洲分布的大花菟丝子 (*Cuscuta reflexa*) 叶绿体基因组核苷酸全序列进行测定, 并与已经发表的分布于欧洲的大花菟丝子进行了比较分析。研究结果表明亚洲分布的大花菟丝子叶绿体基因组总长度为 120 972 bp, 由 79 499 bp 的长单拷贝区, 8 369 bp 的短单拷贝区, 以及两个 16 552 bp 的反向重复区组成, 其长度比欧洲分布的大花菟丝子小了 549 bp; 基因组总 GC 含量为 38.3%, 稍高于欧洲分布的大花菟丝子。两地区的大花菟丝子叶绿体全基因组编码的功能基因完全相同, 且基因排列顺序也完全一致。另外, 经过进一步序列比对后发现亚洲分布的大花菟丝子与欧洲分布的存在 251 个插入和 210 个缺失现象, 总插入缺失及碱基替换长度分别为 7 649 bp 和 3 720 bp, 最大的插入和缺失长度分别为 426 bp 和 435 bp。很多插入缺失都是单碱基, 但仍然存在四个长度超过 200 bp 的大突变, 两个大的缺失发生在 *ycf2* 基因中, 两个大的插入分别发生在 *trn*F-*psb*E 和 *mat*K-*trn*Q 间隔区, 详细的对比后发现大量的插入缺失都发生在大单拷贝区的基因间隔区, 且插入缺失在反向重复区的发生频率较低。本研究首次报道不同大洲分布的同种异养植物的叶绿体全基因组比较分析, 为研究这两个区域的居群多样性提供了基础资料。

关键词: 大花菟丝子; 叶绿体基因组; 插入缺失; 对比分析
中图分类号: Q 75 **文献标识码**: A **文章编号**: 2095-0845(2013)02-158-07

Comparative Analyses of Indels Based on the Whole Chloroplast Genome of *Cuscuta reflexa* between European and Asian Populations

WANG Chao-Bo[1,2], GONG Xun[1**]
(1 *Key Laboratory of Economic Plant and Biotechnology*, *Kunming Institute of Botany*, *Chinese Academy of Sciences*, Kunming 650201, China; 2 *University of Chinese Academy of Sciences*, Beijing 100049, China)

Abstract: We determined the complete nucleotide sequence of the chloroplast genome of Asian *Cuscuta reflexa* which is a world distributed dodder using shotgun method at Illumina's Genome Analyzer, and then compared it with the corresponding published sequence of the same species that distributed in Europe. The chloroplast genome length was 120 972 bp, which was 549 bp shorter than the European one, with a large single copy (LSC) region of 79 499 bp, a small single copy (SSC) region of 8 369 bp and two inverted repeats (IR) of 16 552 bp each. The overall GC content was 38.3%, which is slightly higher than European *C. reflexa*. The chloroplast genome of *C. reflexa* from both area encoded identical functional genes in the same order. On the other hand, detailed analyses revealed 251 insertions and 210 deletions. The length of indels substitution sum to 7 649 bp and the largest insertion and deletion reached 426 bp and 435 bp respectively. Meanwhile, 3 720 bp base substitution events were found in the entire chloroplast genome of Asian *C. reflaxa*. Majority of the indels observed were single-base but four large length mutations

* Foundation item: The Joint Funds of the National Natural Science Foundation of China and Yunnan Provincial Government (U1136602)
** Author for correspondence; E-mail: gongxun@mail.kib.ac.cn
Received date: 2012-05-07, Accepted date: 2012-06-15
作者简介: 王朝波 (1984-) 男, 硕士, 主要从事异养植物叶绿体基因组研究。

longer than 200 bp were also detected, including two deletions in *ycf2* region, one insertion in *trn*F-*psb*E and another insertion in *mat*K-*trn*Q. Most indels located in the intergenic regions of the LSC region while rare in the IR region. This research initially compared the whole chloroplast genome of same heterotrophic plants distributed in different continents.

Key words: *Cuscuta reflexa*; Chloroplast genome; Indels; Comparative analysis

The chloroplast genome of photosynthetic plants consists of multiple copies of homogeneous circular double-stranded DNA molecules ranging in size from 120 to 217 kb (Kumar *et al.*, 2009; Wicke *et al.*, 2011), while the non-photosynthetic (heterotrophic) plants consists of a genome size less than 120 kb in general, such as holoparasitic plant *Epifagus virginiana* (Wolfe *et al.*, 1992), *Cuscuta gronovii*, *C. obtusiflora*, hemiparasitic plant *C. reflexa*, *C. exaltata* (Funk *et al.*, 2007; McNeal *et al.*, 2007), and myco-heterotrophic plant *Neottia nidus-avis* (Logacheva *et al.*, 2011), *Rhizanthella gardneri* (Delannoy *et al.*, 2011). Whole chloroplast genome of heterotrophic plant has lost all or part of their photosynthetic genes. Besides, the overall structure of chloroplast genome is generally conserved despite some mutations observed (Krause, 2011) including small structural changes such as inversions (Saski *et al.*, 2005), translocations (Lee *et al.*, 2007; Hirao *et al.*, 2008), indels (Asano *et al.*, 2004; Shahid Masood *et al.*, 2004; Chung *et al.*, 2006; Jo *et al.*, 2011), base substitutions (Morton and Clegg, 1995), and extant rearrangement (Haberle *et al.*, 2008; Guisinger *et al.*, 2011) etc. Three inversions (*rpo*B-*psb*D, *trn*T-*pet*L, *ccs*A-*trn*L) has been found in *C. reflexa* compared to the classic structure of chloroplast genome (Funk *et al.*, 2007).

It is reported that indels do not occur at random locations within organelle genomes, but they are often associated with specific DNA sequence features. Regions containing repeats that lead to slipped strand mispairing and intramolecular recombination are thought to cause the majority of indel mutations (reviewed by (Kelchner, 2000)). Length mutations have been reported in genus *Saccharum* (Asano *et al.*, 2004) and *Solanum* (Chung *et al.*, 2006) and so on. In the intergenic region of this tRNA gene cluster, a large number of indels ranging from 1 to 811 bp in length are commonly presented. The hotspot of divergence is also recognized in the region associated with tRNAs and the largest deletions over 500 bp in size are found at upstream of *trn*C (GCA) and *trn*T (GGU), and at downstream of *trn*D (GUC) (Maier *et al.*, 1995; Ogihara *et al.*, 2002; Calsa Júnior *et al.*, 2004). Another hotspot of divergence is in the *ycf1* and *ycf2* gene which is conserved and located along with *ycf15* (Plader *et al.*, 2007; Haberle *et al.*, 2008).

Here, the entire chloroplast genome nucleotide sequence of the *C. reflexa* distributed in Asia was reported, which may interpret the worldwide spreading of this parasitic plant. We especially focus on the length mutations and indels located in coding region. Our study will provide a rich source of the nucleotide and amino acid sequence data, which can be utilized to address phylogenetic and molecular evolutionary question as well as to heterotrophic biology.

Materials and methods

The material of *C. reflexa* was collected from Dong'e Town (23.6137879°N, 102.015824°E), Yuanjiang County, Yunnan Province, China. The chloroplast DNA was isolated from 100 g fresh leaves using an improved extraction method that includes high ionic strength buffer at low pH (3.6) buffer (Triboush *et al.*, 1998). Chloroplast DNA was sequenced with Illumina's Genome Analyzer at Beijing Genomics Institute (BGI) in Shenzhen, China. Subsequently, SOAPdenovo were used to assemble the sequence reads of chloroplast genomes (Li *et al.*, 2009). Eight small gaps were closed by PCR method using Qiagen Taq Polymerase. Primers used for PCR

were designed using available information from both ends of the gap. All primers were ordered from Sangon Biotech (Shanghai, China) as shown in Table 1. PCR products were analyzed by electrophoresis on 1% agarose gel and then sequenced using standard Sanger protocols on ABI 3730 instruments, and the thermal cycling program was as follows: 80 ℃ for 5 min; then 33 cycles of [95 ℃ for 45 s; 42–52 ℃ (depending upon the annealing temperature) for 45 s; 65 ℃, 1 min], and 65 ℃ for 5 min.

Annotation of the sequenced genome was performed using DOGMA (Wyman *et al.*, 2004), complement with adjustment for start and stop codons and for intron/exon boundaries manually. Gene map was drawn using OGDRAW 1.2 (Lohse *et al.*, 2007), and each indels was mapped onto the exact position of the map. Alignments of two *C. reflexa* chloroplast genomes were performed using MAFFT version 6 (Katoh and Toh, 2008). Indels were calculated using DnaSP v5 (Librado and Rozas, 2009).

Result and discussion

The entire chloroplast genome is a circular double-stranded DNA molecule of 120 972 bp, which is shorter than the European one by 549 bp (Funk *et al.*, 2007). It exhibited a quadripartite structure, similar to typical angiosperm chloroplast genomes, with a large single copy region of 79 499 bp, a small single copy region of 8 369 bp and two inverted repeats of each 16 552 bp containing four rRNA operons. The overall GC content was 38.3%, which is slightly higher than European *C. reflexa* (38.2%) (Funk *et al.*, 2007). *C. reflexa* chloroplast genome in both area encoded identical functional genes with same gene order. A total of 116 genes were identified, 15 (including six tRNAs and four rRNAs) of which were duplicated in IR. Ten genes contained one intron and three of them contained two introns (*rps12*, *ycf3*, *clpP*), and four of the 13 genes with introns were tRNAs. All the *ndh* genes (except *ndhB*), *infA*, *trnK-UUU* was completely loss from the chloroplast genome. Two ribosomal protein genes (*rpl23*, *rps16*), two tRNA genes (*trnV-UAC*, *trnG-UCC*), redundant partial of *ycf2* which repeated in the IR, together with *ycf15* and *ndh*B remained as pseudo genes.

The main purpose of this study was to evaluate the genetic diversity between two *C. reflexa* using entire chloroplast DNA sequence. Comparative sequence analysis revealed a large number of mutations throughout the genome that includes indels and base substitutions. In total, there were 7 649 bp Indels (consisting of 251 insertions, 210 deletions), and

Table 1 Primers designed for closing gaps

No.	Name	Position	Length	Primers (5′—3′)	Region
1	ca1f	4963	589	TGTTGCGCTCTTCATCTTT	*psb*I
	ca1r	5552		TACTCGCTGCTACAATCCA	*trn*S-GCU
2	ca2f	18230	451	AAATCGGACGTGAATGTTT	*rpo*C1 intron
	ca2r	18681		ATTATGTTCCCGTAAGCAA	*rpo*C1 intron
3	ca3f	29256	1016	TATCTAATGCGTTCTCCCA	*psb*C
	ca3r	30272		TGTTACTTGACCAGCCCTC	*trn*S-UGA
4	ca4f	36531	1175	GAGGCATTCCCGTATCTAA	*psa*A
	ca4r	37706		AACGAAATCCATTCTTACCA	*ycf3*
5	ca5f	38616	1180	ACTTGGCGTGTCTGTCTTT	*ycf3*
	ca5r	39796		AAATGGGTCGGTTTGAAGA	*trn*S-GGA
6	ca6f	40752	969	ACCCATAGAGTTGGAAGTG	*trn*T-UGU
	ca6r	41721		ACAGGATTTGGCTCAGGAT	*trn*L-UAA
7	ca7f	43358	1149	CGATGGTTGGCTGTTCACG	*psb*F
	ca7r	44507		GCGGATTTGGTCAGGGAGA	*pet*A
8	ca8f	74330	980	CTCTATGCCTTGCGGTAAT	*rpl2*
	ca8r	75310		TTTGGTCCACGAATCTAAT	*ycf2*

3 720 bp base substitution events in the entire chloroplast genome of Asian *C. reflaxa*. Overall genome wide indel events in Asian population relative to European population are shown in Fig. 1. The Largest insertion and deletion reached 426 bp and 435 bp respectively (Fig. 2).

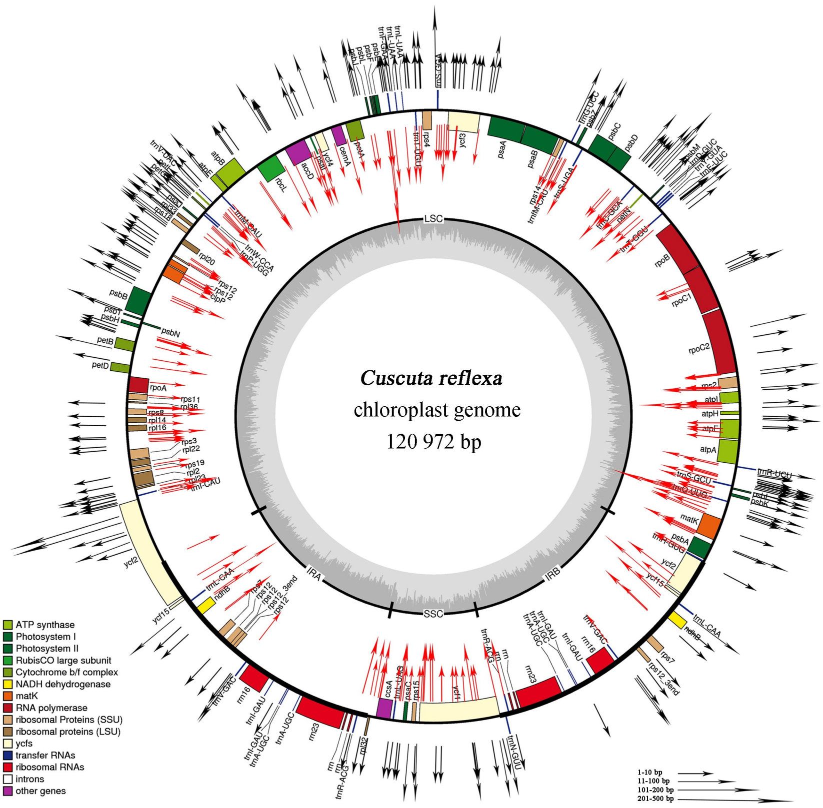

Fig. 1　Comparative sequence analysis of the two *C. reflexa* chloroplast genomes. Gene order and content of Asian *C. reflexa* are the same with the European *C. reflexa* chloroplast genomes. Genes shown inside the circle are transcribed clockwise and those outside are transcribed counterclockwise. Dashed area of the inner circle drew the GC content of the chloroplast genome. Red arrows indicate deletion event in Asian *C. reflexa*, and black arrows indicate insertion event in Asian *C. reflexa*. Length of arrow indicates the size of indel event

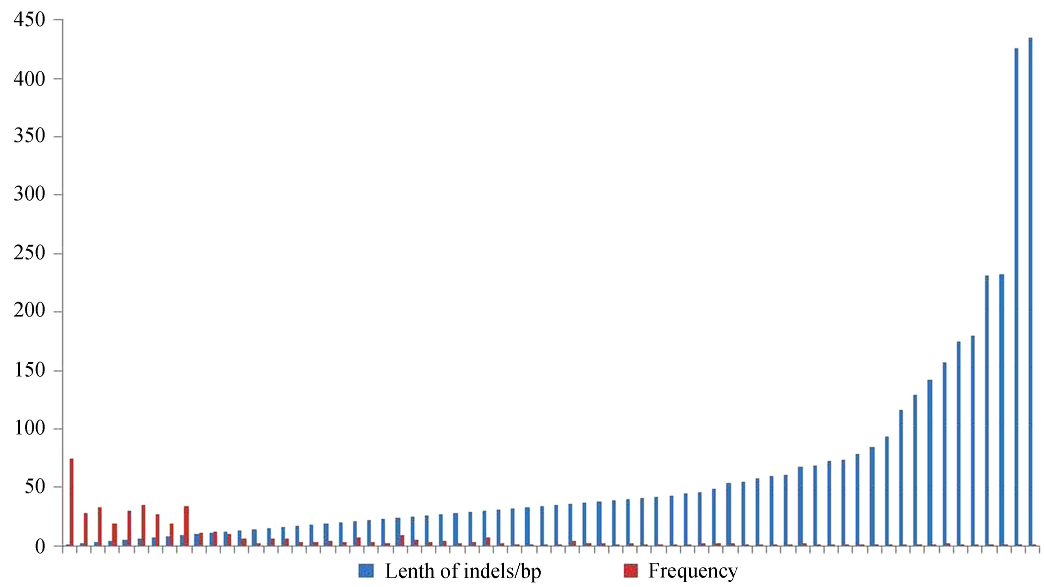

Fig. 2　The length of each indel scored as a character for two *C. reflexa*. Frequency is the observed number of indel events with certain length

Of the 251 insertion events, only 28 (11.2%) were identified in the coding region, 47 (18.7%) in intron and 176 (70.1%) in the intergenic region (Table 2). Most of the insertions (49) were single-base, however multi-nucleotide insertions were also identified. As shown in Table 3, most of the insertions that located in the coding region had a triplex type, which wouldn't disrupt the coding sequence. However, two small insertions of 7-bp and 1-bp were found before the stop codon of *rpl20*, leading to a 42-bp delayed of stop codon. Most insertions in the intron and intergenic region were short, 11 bases or less, and located in the LSC region. The two largest insertions located within the single copy part of *ycf2*, with no influence on the gene function.

The deletion events in three regions (coding, intergenic and intron) were also given in Table 2. Among 210 deletion events, 30 (14.3%) were located in the coding region, 37 in the intron (17.6%) and 143 (68.1%) in the intergenic region. Similar to insertions, numerous single base pair deletions (26) existed in the chloroplast genome of *C. reflaxa*. A single large deletion of 157-bp was located in the coding region of pseudo gene *ycf15*. Since *ycf15* in both area had no function, this deletion was assumed to cause parasitic lifestyle of *C. reflaxa*. Single-base deletion located before stop codon of *pet*A, which made stop codon of this photosynthesis gene 489 bp afterward. The effect of this deletion to the expression of protein is unknown yet. Another 8-bp deletion was found within *rpo*C2 which was also present in the LSC region; resulting in the ahead of stop codon by 12 bp.

Table 2 Indel events in the IR, LSC, and SSC region of *C. reflexa* chloroplast genome

Region Type	IR (Ins/Del)	LSC (Ins/Del)	SSC (Ins/Del)	Total (Percentage) (Ins/Del)
Coding region	2/6	18/12	8/12	28(11.2)/30(14.3)
Intron	2/0	45/37	0/0	47(18.7)/37(17.6)
Intergenic region	19/10	168/143	13/17	176(70.1)/143(68.1)

The values in parentheses represent proportion (%). IR = inverted repeat, LSC = large single copy, SSC = small single copy. Ins = Number of insertion, Del = Number of deletion

Table 3 Counted number and length of indels located in coding region

Gene	Ins	Lenth/bp	Dels	Lenth/bp
*mat*K	1	3	2	24, 36
*rpo*C2	3	9, 6, 21	1	**8**
*trn*S-GGA	0		1	3
*trn*G-UCC	1	**16**	0	
*pet*A	0		1	**1**
*acc*D	0		2	9, 54
*atp*E	1	6	0	
*trn*V-UAC	0		1	**37**
rps18	1	12	0	
rpl20	2	**1, 7**	0	
*rpo*A	0		1	9
rps8	0		1	21
rpl2	1	6	0	
rpl23	2	6, 15	1	**8**
ycf2	6		3 (2)	9, 24, 9 (24, 9)
ycf15	0		1 (1)	**157 (157)**
ycf1	8	9, 9, 6, 12, 6, 3, 21, 3	12	36, 12, 12, 24, 9, 9, 6, 6, 9, 42, 6, 9
*ndh*B	1 (1)	**2 (2)**	0	

The values in parentheses represent indels repeated in the IR region. The values in bold represent indels that does not happened in triplex type. Ins = insertions, Dels = deletions

Overall, the indels were observed more frequently in the intergenic region rather than intron and coding region. It is known that direct repeats contribute significantly to variability of chloroplast genomes (Wolfe *et al.*, 1992; Ogihara *et al.*, 2002). Shashid *et al.* (2004) and Hiratsuka *et al.* (1989) also identified many indels in the intergenic region when comparing rice and tobacco. Our research showed that heterotrophic plant undergone a much relaxed selective constraint compared to autotrophic plant, especially in the intergenic region.

A total of 39 (only 8.5%) indel events happened within IR region, suggesting that IR region contributed greatly to the conservation of nucleotide sequence (Wicke *et al.*, 2011). The complete a-

lignment of chloroplast genome sequences revealed a reasonable number of indels in the intron and intergenic spacer regions in *C. reflexa*, but there were fewer indels occurred in 18 genes (Table 3) when compared with Eropean *C. reflexa*. Comparison of *acc*D genes between two *C. reflexa* is similar to that of seven Solanaceae species, two deletion events of 9 and 54 bp were observed. The *acc*D gene encodes the beta-carboxyl transferase subunits of acetyl-CoA-carboxylase (ACCase) and is present in the plastids of most land plants. The tobacco ACCase is essential for leaf development, leaf longevity, and seed yield (Madoka *et al.*, 2002). *C. reflexa* lost leaves through their entire lifecycle, we still observed the relatively long deletion (54 bp) which kept accD as a functional gene, showed that accD gene maybe important at seedling stage. Additional indels did take place in *ycf1* and *ycf2*. Most vascular plants contain these two genes, which appear to be essential for cell survival, and often seemed to be housekeeping genes (Krause, 2011). A total of 20 indels occurred within *ycf1* and 11 within *ycf2*; contributing to 53.4% of the indels within coding region. Our observation consistent with former report that these two genes varied greatly among land plants (Moore *et al.*, 2010).

As for the intergenic and intron region, most multi-nucleotide deletions were 2 to 11 bp. Seven remarkable mutation regions identified including *trn*M-*trn*V, *trn*S-*trn*R, *rpl*16-*rps*3, *clp*P-*psb*B, *trn*C-*psb*D, *mat*K-*trn*Q, and *ycf*3 introns. Striking feature of *C. reflexa* was two largest deletions, of which 232-bp deletion was in *mat*K-*trn*Q and a 426-bp deletion in *trn*F-*psb*E.

A coarse comparison of our results with those of the Poaceae plants showed that the length mutation was bounded with the tRNA cluster. In addition, the indel events in *C. reflexa* is about 3 times more than that between wild and cultivated rice (Matsuoka *et al.*, 2002), suggesting that parasitic plants undergone relaxed selective constraint on chloroplast genome along evolutionary history.

In conclusion, the Asian *C. reflexa* chloroplast genome is very similar to European one even though a great number of insertion/deletion and base substitution events were detected. A substantial amount of polymorphism, mainly resulting from indels, was observed within in *C. reflexa* of two areas. Furthermore, the indel events in the intergenic region were quite frequent comparing with other plant. Taken together, our data obtained from Asian *C. reflexa* provide potential molecular markers to evaluate genetic diversity as well as evolutionary processes in heterotrophic plants.

Acknowledgements: We are grateful to Yu Jiaojun and Guan Mengmeng for critically reading the manuscript. The skillful technical assistance of Dr. Ma Pengfei is gratefully acknowledged. We are thankful to lab members for their generous help.

References:

Asano T, Tsudzuki T, Takahashi S *et al.*, 2004. Complete nucleotide sequence of the sugarcane (*Saccharum officinarum*) chloroplast genome: a comparative analysis of four monocot chloroplast genomes [J]. *DNA Research*, **11** (2): 93—99

Calsa Júnior T, Carraro DM, Benatti MR *et al.*, 2004. Structural features and transcript-editing analysis of sugarcane (*Saccharum officinarum* L.) chloroplast genome [J]. *Current Genetics*, **46** (6): 366—373

Chung HJ, Jung JD, Park HW *et al.*, 2006. The complete chloroplast genome sequences of *Solanum tuberosum* and comparative analysis with Solanaceae species identified the presence of a 241-bp deletion in cultivated potato chloroplast DNA sequence [J]. *Plant Cell Reports*, **25** (12): 1369—1379

Delannoy E, Fujii S, Colas des Francs-Small C *et al.*, 2011. Rampant gene loss in the underground orchid *Rhizanthella gardneri* highlights evolutionary constraints on plastid genomes [J]. *Molecular Biology and Evolution*, **28** (7): 2077—2086

Funk HT, Berg S, Krupinska K *et al.*, 2007. Complete DNA sequences of the plastid genomes of two parasitic flowering plant species, *Cuscuta reflexa* and *Cuscuta gronovii* [J]. *BMC Plant Biology*, **7** (45): 1—12

Guisinger MM, Kuehl JV, Boore JL *et al.*, 2011. Extreme reconfiguration of plastid genomes in the angiosperm family Geraniaceae: rearrangements, repeats, and codon usage [J]. *Molecular Biology and Evolution*, **28** (1): 583—600

Haberle RC, Fourcade HM, Boore JL *et al.*, 2008. Extensive rearrangements in the chloroplast genome of *Trachelium caeruleum* are associated with repeats and tRNA genes [J]. *Journal of Mo-*

lecular Evolution, **66** (4): 350—361

Hirao T, Watanabe A, Kurita M *et al.*, 2008. Complete nucleotide sequence of the *Cryptomeria japonica* D. Don. chloroplast genome and comparative chloroplast genomics: diversified genomic structure of coniferous species [J]. *BMC Plant Biology*, **8** (70): 1—20

Hiratsuka J, Shimada H, Whittier R *et al.*, 1989. The complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals [J]. *Molecular and General Genetics*, **217** (2-3): 185—194

Jo YD, Park J, Kim J *et al.*, 2011. Complete sequencing and comparative analyses of the pepper (*Capsicum annuum* L.) plastome revealed high frequency of tandem repeats and large insertion/deletions on pepper plastome [J]. *Plant Cell Reports*, **30** (2): 217—229

Katoh K, Toh H, 2008. Recent developments in the MAFFT multiple sequence alignment program [J]. *Briefings in Bioinformatics*, **9** (4): 286—298

Kelchner SA, 2000. The evolution of non-coding chloroplast DNA and its application in plant systematics [J]. *Annals of The Missouri Botanical Garden*, **87** (4): 482—498

Krause K, 2011. Piecing together the puzzle of parasitic plant plastome evolution [J]. *Planta*, **234** (4): 647—656

Kumar S, Hahn FM, McMahan CM *et al.*, 2009. Comparative analysis of the complete sequence of the plastid genome of *Parthenium argentatum* and identification of DNA barcodes to differentiate *Parthenium* species and lines [J]. *BMC Plant Biology*, **9** (131): 1—12

Lee HL, Jansen RK, Chumley TW *et al.*, 2007. Gene relocations within chloroplast genomes of *Jasminum* and *Menodora* (Oleaceae) are due to multiple, overlapping inversions [J]. *Molecular Biology and Evolution*, **24** (5-6): 1161—1180

Li RQ, Yu C, Li YR *et al.*, 2009. SOAP2: an improved ultrafast tool for short read alignment [J]. *Bioinformatics*, **25** (15): 1966—1967

Librado P, Rozas J, 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data [J]. *Bioinformatics*, **25** (11): 1451—1452

Logacheva MD, Schelkunov MI, Penin AA, 2011. Sequencing and analysis of plastid genome in mycoheterotrophic orchid *Neottia nidus-avis* [J]. *Genome Biology and Evolution*, **3** (1): 1296—1303

Lohse M, Drechsel O, Bock R, 2007. OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes [J]. *Current Genetics*, **52** (5-6): 267—274

Madoka Y, Tomizawa K, Mizoi J *et al.*, 2002. Chloroplast transformation with modified accD operon increases acetyl-CoA carboxylase and causes extension of leaf longevity and increase in seed yield in tobacco [J]. *Plant and Cell Physiology*, **43** (12): 1518—1525

Maier RM, Neckermann K, Igloi GL *et al.*, 1995. Complete sequence of the maize chloroplast genome: gene content, hotspots of divergence and fine tuning of genetic information by transcript editing [J]. *Journal of Molecular Biology*, **251** (5): 614—628

Matsuoka Y, Yamazaki Y, Ogihara Y *et al.*, 2002. Whole chloroplast genome comparison of rice, maize, and wheat: implications for chloroplast gene diversification and phylogeny of cereals [J]. *Molecular Biology and Evolution*, **19** (12): 2084—2091

McNeal JR, Kuehl JV, Boore JL *et al.*, 2007. Complete plastid genome sequences suggest strong selection for retention of photosynthetic genes in the parasitic plant genus *Cuscuta* [J]. *BMC Plant Biology*, **7** (57): 1—22

Moore MJ, Soltis PS, Bell CD *et al.*, 2010. Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots [J]. *Proceedings of the National Academy of Sciences*, **107** (10): 4623—4628

Ogihara Y, Isono K, Kojima T *et al.*, 2002. Structural features of a wheat plastome as revealed by complete sequencing of chloroplast DNA [J]. *Molecular Genetics and Genomics*, **266** (5): 740—746

Plader W, Yukawa Y, Sugiura M *et al.*, 2007. The complete structure of the cucumber (*Cucumis sativus* L.) chloroplast genome: its composition and comparative analysis [J]. *Cellular & Molecular Biology Letters*, **12** (4): 584—594

Saski C, Lee SB, Daniell H *et al.*, 2005. Complete chloroplast genome sequence of *Gycine max* and comparative analyses with other legume genomes [J]. *Plant Molecular Biology*, **59** (2): 309—322

Shahid MM, Nishikawa T, Fukuoka S *et al.*, 2004. The complete nucleotide sequence of wild rice (*Oryza nivara*) chloroplast genome: first genome wide comparative sequence analysis of wild and cultivated rice [J]. *Gene*, **340** (1): 133—139

Triboush SO, Danilenko NG, Davydenko OG, 1998. A method for isolation of chloroplast DNA and mitochondrial DNA from sunflower [J]. *Plant Molecular Biology*, **16** (2): 183—183

Wicke S, Schneeweiss GM, dePamphilis CW *et al.*, 2011. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function [J]. *Plant Molecular Biology*, **76** (3-5): 273—297

Wolfe KH, Morden CW, Palmer JD, 1992. Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant [J]. *Proceedings of the National Academy of Sciences*, **89** (22): 10648—10652

Wyman S, Jansen R, Boore J, 2004. Automatic annotation of organellar genomes with DOGMA [J]. *Bioinformatics*, **20** (1): 3252—3255